Global model comparison: DIMOSIC
Different models – same initial conditions

WGNE Session 2021

Linus Magnusson and Ron Mctaggart-Cowan

Simon Lang (ECMWF)

Martin Köhler (DWD)

Yves Bouteloup (MeteoFrance)

Duncan Ackerley (UKMO)

Ron Mctaggart-Cowan (Environment Canada)

Jan-Huey Chen, Linjiong Zhuo (GFDL)

James Doyle, Alex Reinecke (NRL)

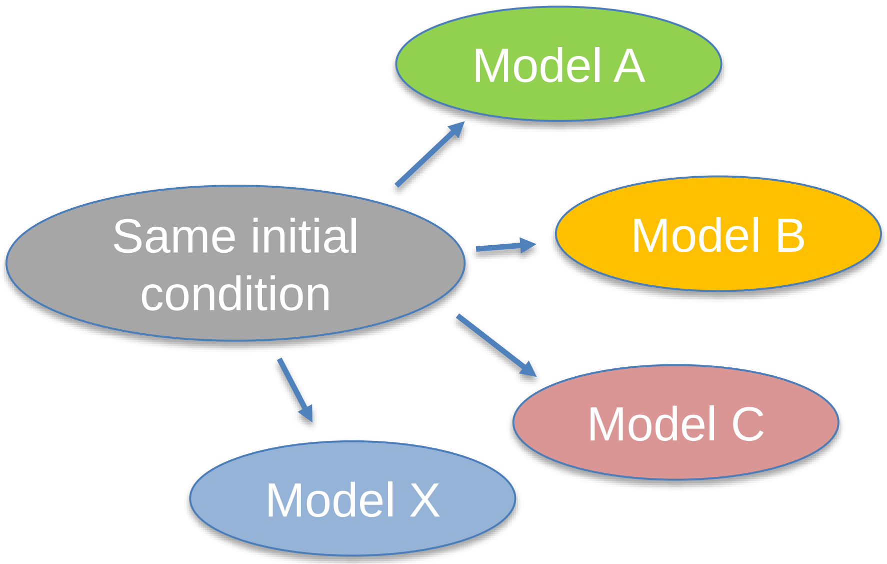
Munehiko Yamaguchi (JMA), Takumi Matsunobu (U. Tsukuba, LMU), Mio Matsueda (U. Tsukuba)

Young Cheol Kwon, Junsu Kim (KMA)

Julian Quinting (KIT)

Global model comparison: DIMOSIC

Different models – same initial conditions



Run several models from the same initial conditions for a set of dates covering one year

Model resolutions ~10-25 km
Initialised from ECMWF initial conditions
2018-06-06 to 2019-06-06, 00UTC, every 3rd day

Current status

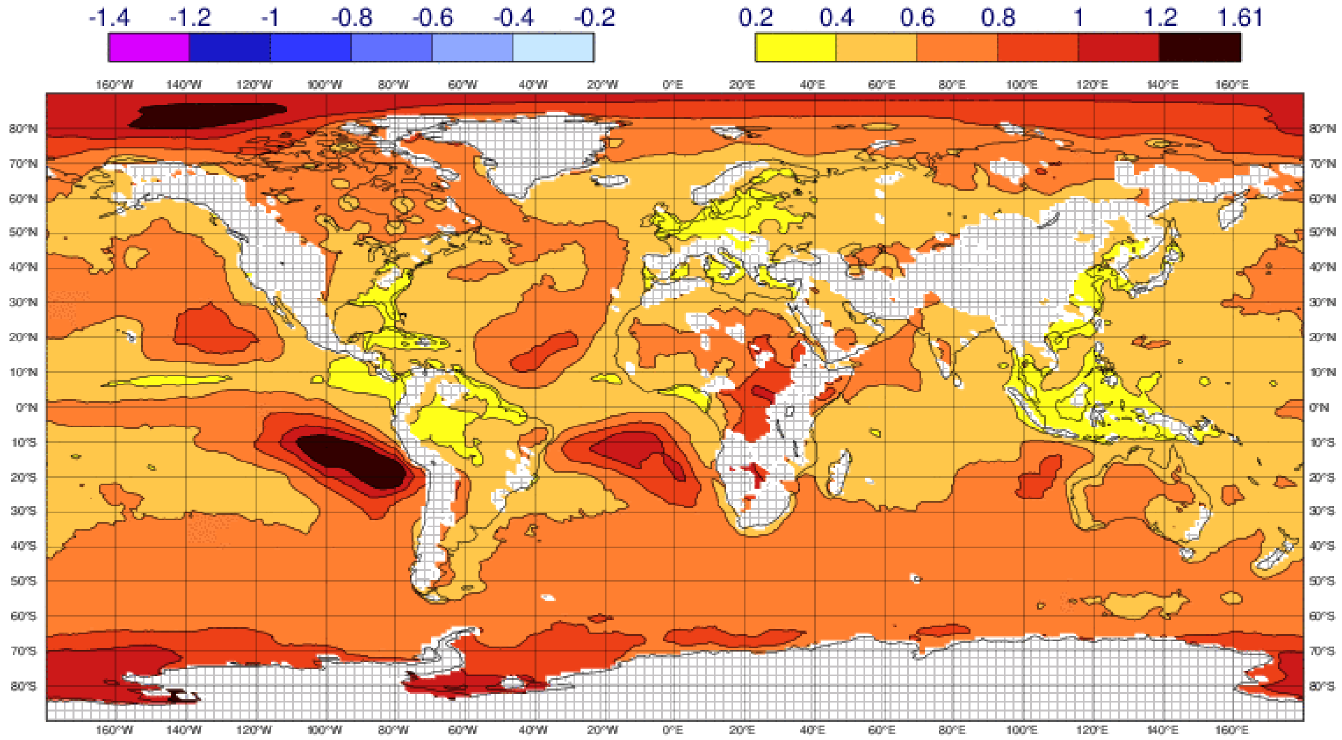
- ECMWF/IFS – 9km operations / 47r1 / 47r3
- UK Metoffice/UM – 10km ready
- DWD/ICON - 13km ready (updated in April 21)
- MeteoFrance/ARPEGE – 5/25km - ready + New physics version
- GFDL/SHiELD* – 13km ready
- KMA/KIM – 25km ready / working on update
- CMC – 15km ready
- JMA/GSM1705 – 20 km ready
- NRL/Neptune – ongoing tests

* For description of differences between GFS and SHiELD, see https://doi.org/10.1029/2020MS002223

Data Preparation and Evaluation Reference

Multi-analysis for verification based on analyses from TIGGE: ECMWF, UKMO, NCEP, CMC, KMA, JMA

T850 ensemble spread (not mean corrected)



Avoid biasing scores towards ECMWF system.

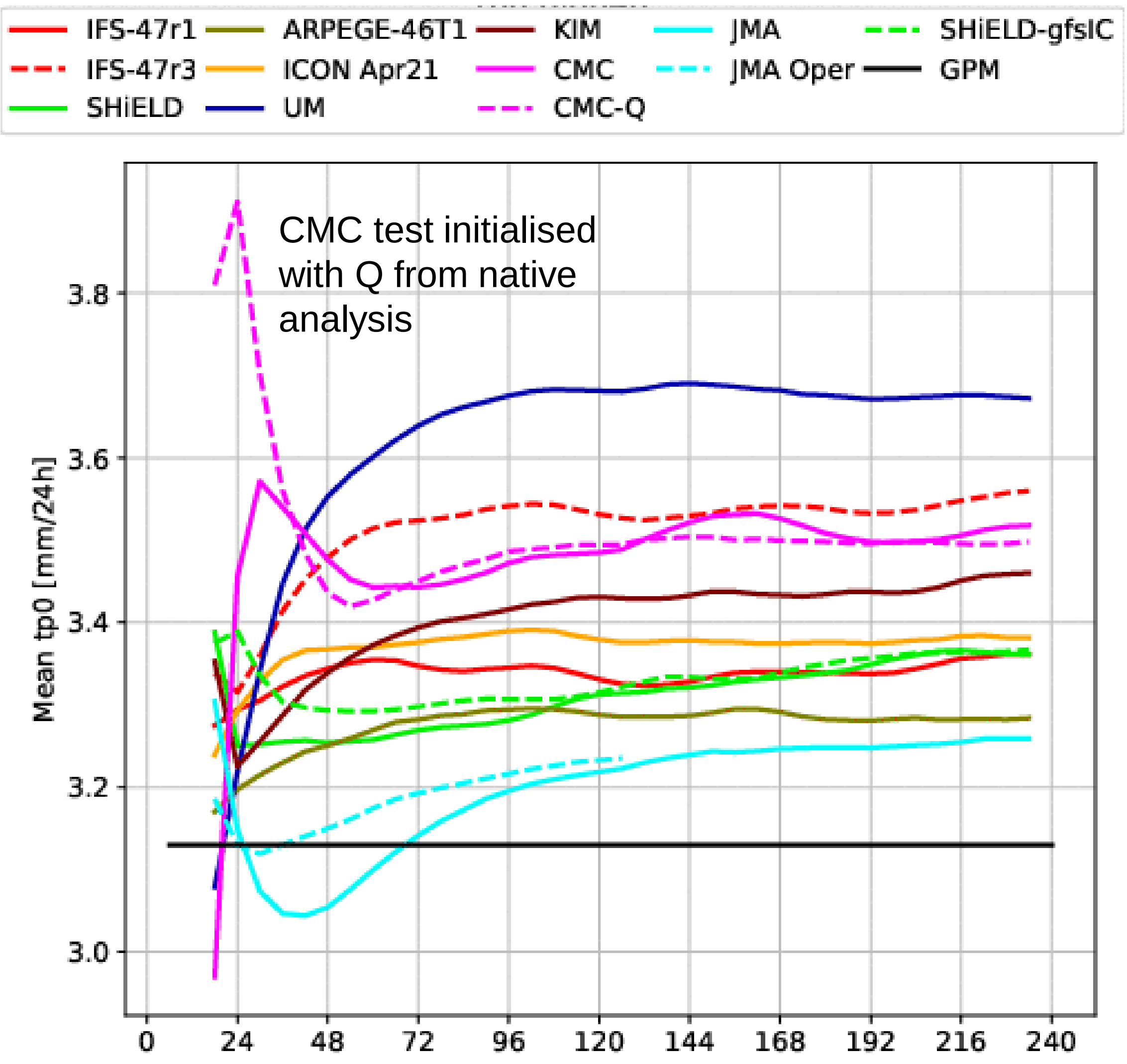
Differences between analyses are maximized in the oceanic stratocumulus regions and at high latitudes.

- Orographic mask based on lowest geopotential and highest orography in all models
- All data evaluated on a common 0.5 degree grid
- Z, T, U, V, (Q) on 850, 700, 500, 300, 200 and 100hPa; MSL, 10u, 10v, TP, TTR, and TSR

Model Spin-Up: Mean Precipitation

Most models show a significant spin-up in precipitation on 3-5 day time scales, consistent with long tropical moisture memory.

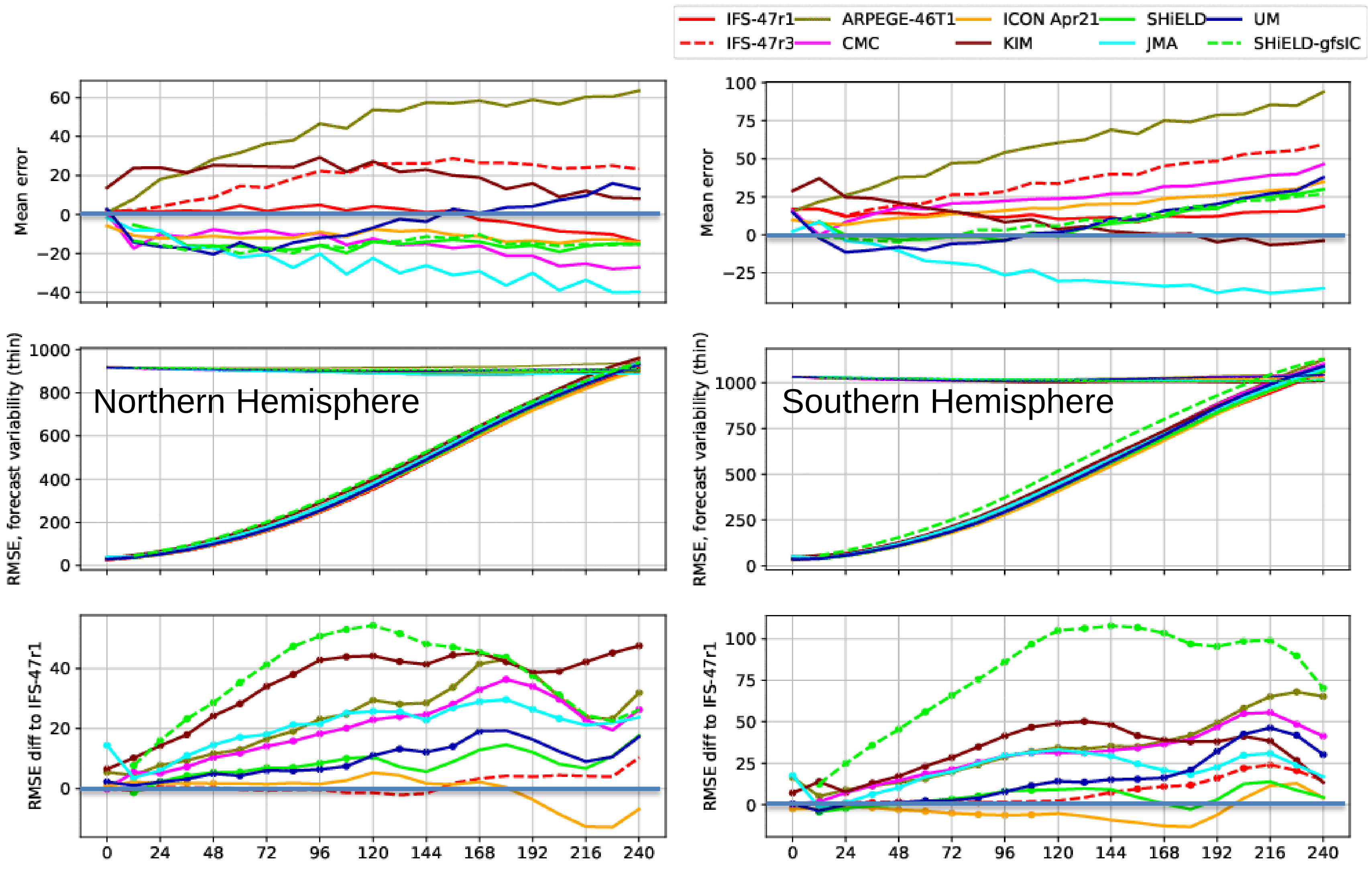
An additional CMC test shows that systematic differences in tropical moisture analyses have a leading-order impact on this behaviour.



24-hour precipitation for running mean windows for 40N-40S

ECMWF

Mid-Tropospheric Height Error Evolution



Broad range of bias and error growth despite the identical initializations.

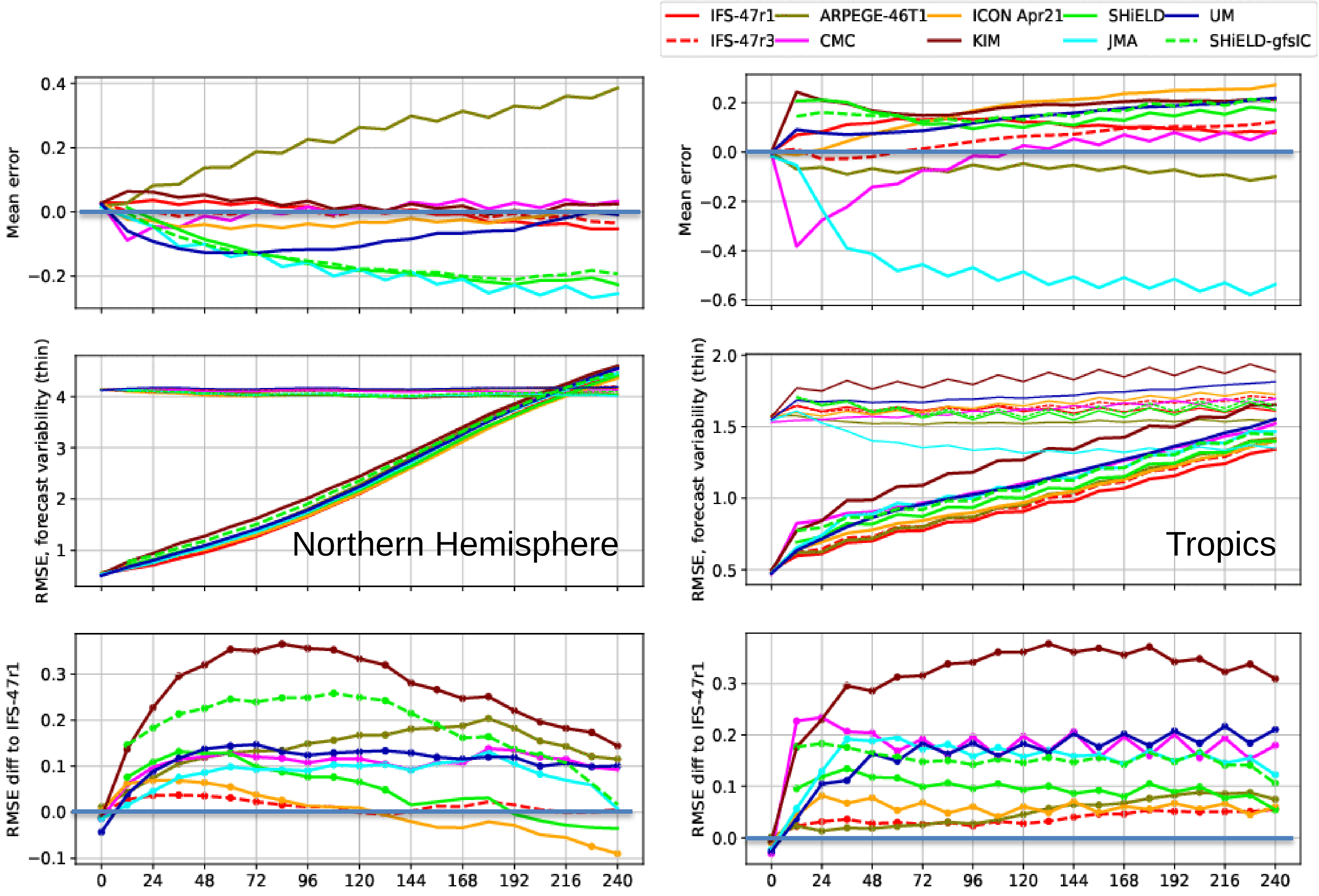
Both systematic and random errors are different across systems.

Evolution of bias (top row), RMSE (middle row) and relative RMSE (w.r.t. IFS; bottom row) of 500 hPa geopotential height in forecast models as shown in the legend for the Northern Hemisphere (left column) and the Southern Hemisphere (right column).

Lower-Tropospheric Temperature Error Evolution

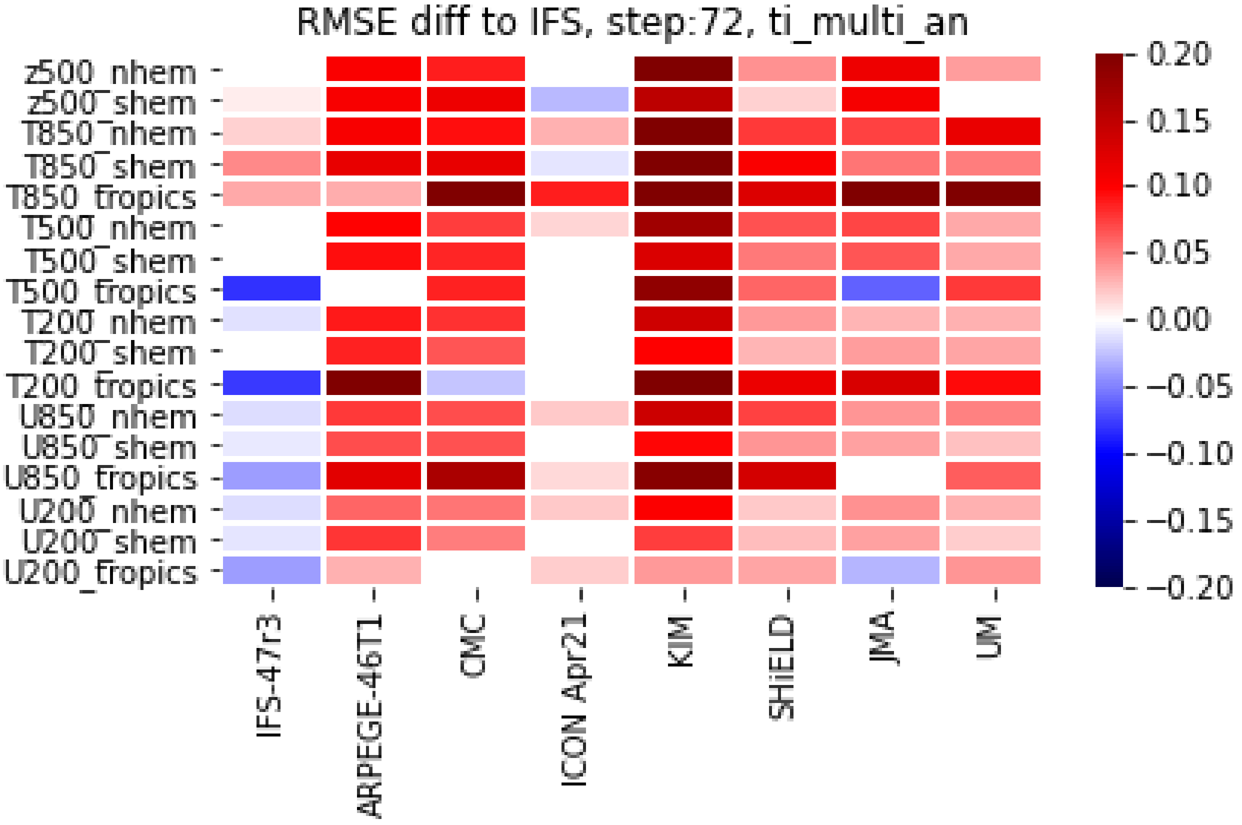
Large biases develop over the first 5 days of some integrations.

Tropical transients may be related to moisture / precipitation spin-up.



Evolution of bias (top row), RMSE (middle row) and relative RMSE (w.r.t. IFS; bottom row) of 850 hPa temperature in forecast models as shown in the legend for the Northern Hemisphere (left column) and the Southern Hemisphere (right column).

Investigating Errors Across Multiple Variables

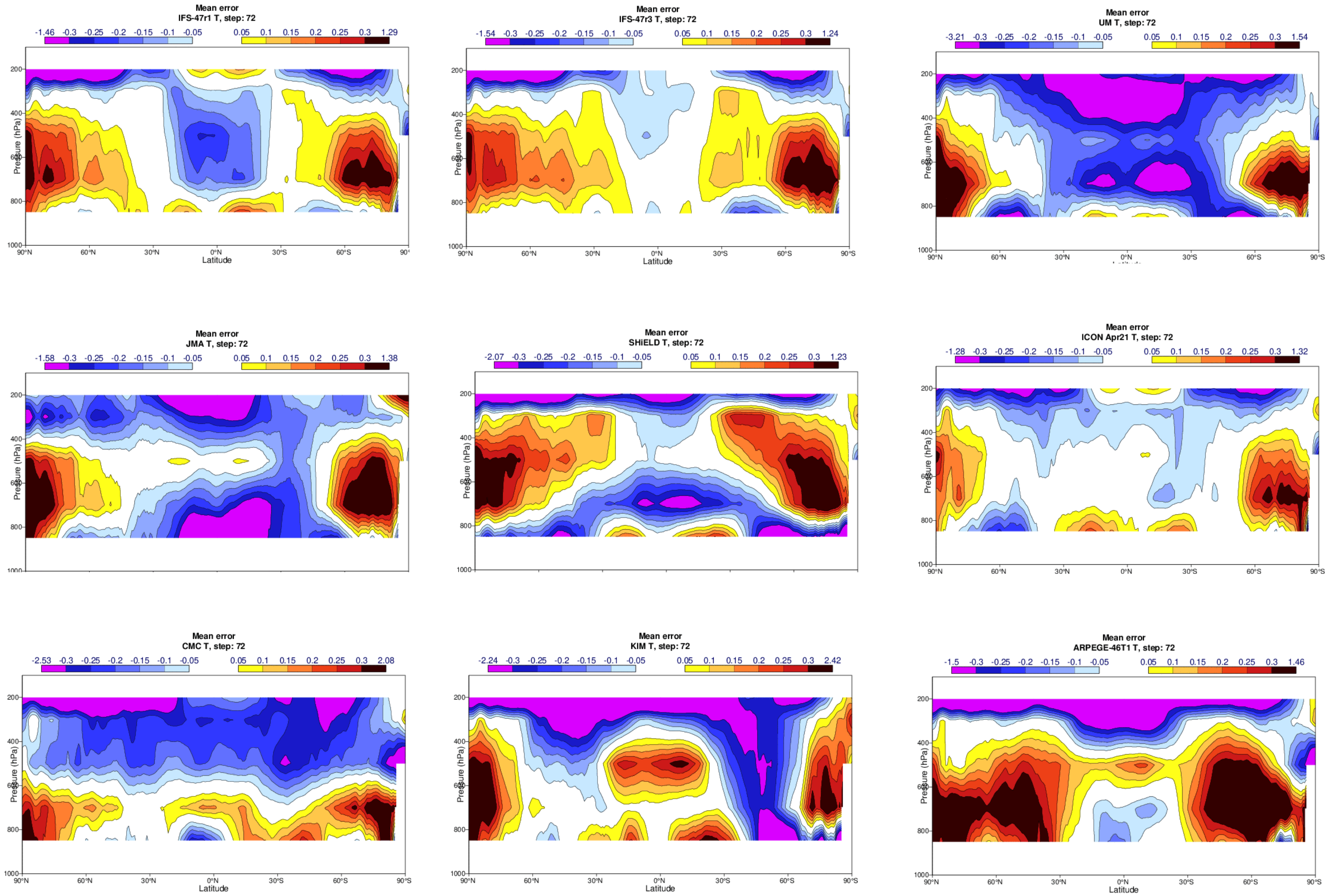


Scorecard for normalized (fractional) difference of RMSE between model (column) and IFS-47r1 for day-3 forecasts of the fields identified on the absicssa. Differences that are not statistically significant are masked with white.

Scorecards help to synthesize the large volume of data to identify both common errors between systems and outliers.

The IFS and ICON models appear to generate the highest quality predictions for the full range of variables investigated.

Temperature Bias Cross-Sections



A more complete error analysis includes bias cross-sections.

Vertical bias structures reflect differences in static stability:

- CMC has reduced mid-tropospheric stability
- Stable PBL in SHIELD
- ICON has smallest biases overall

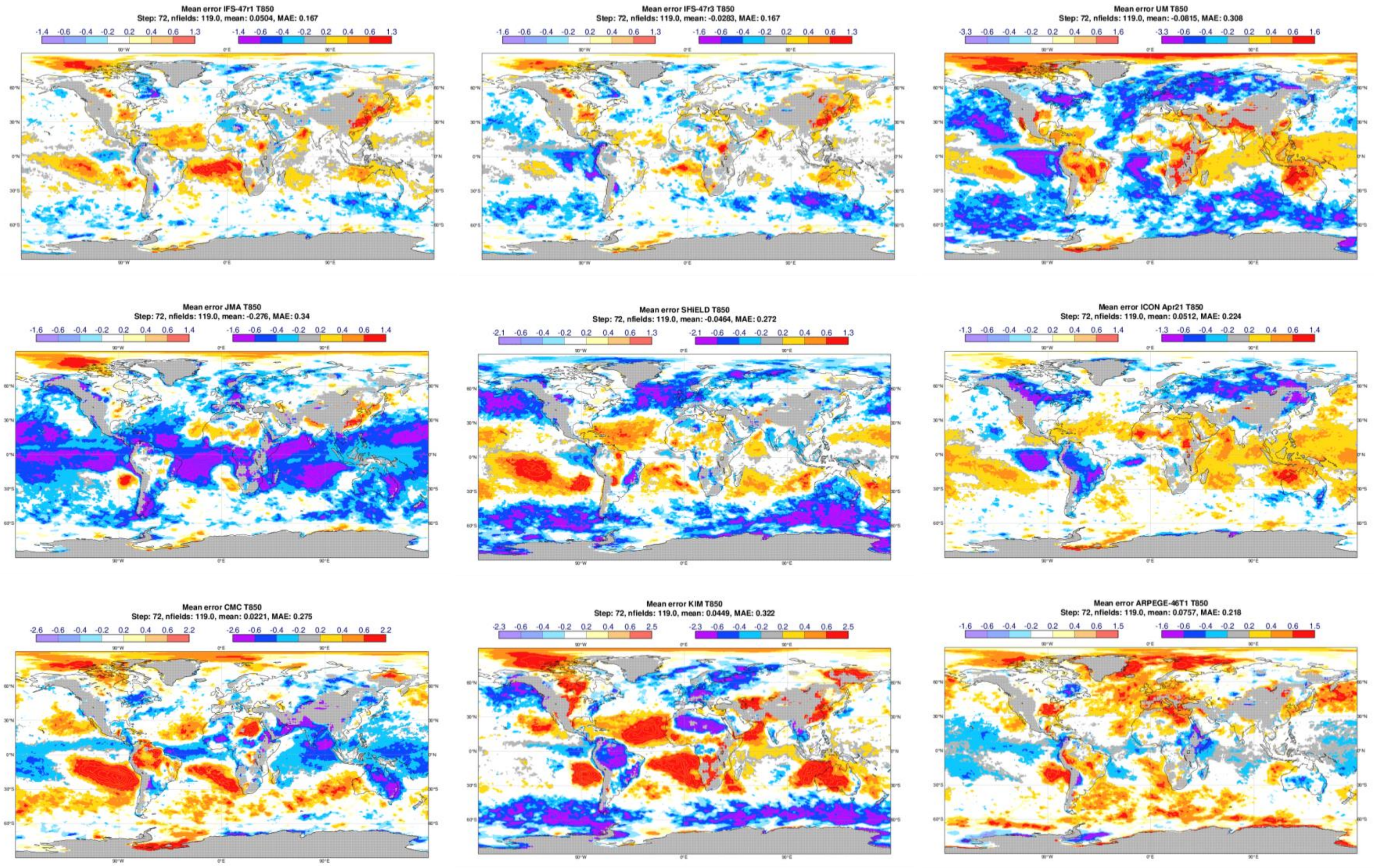
Vertical cross sections of zonal mean day-3 temperature errors (K) in different models as indicated in the panel titles.

Lower-Tropospheric Temperature Biases

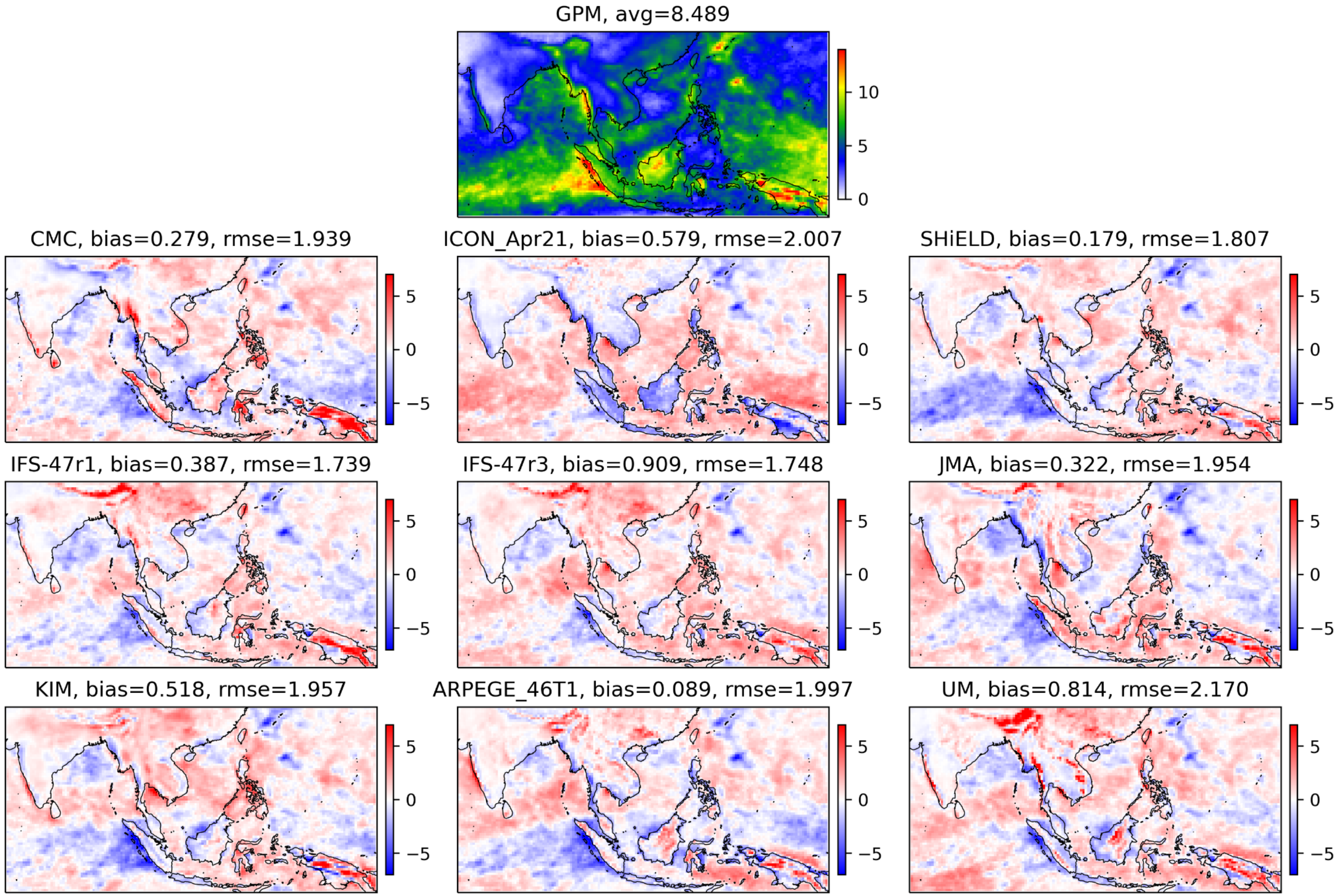
Large spatial variability in biases complicates the interpretation of the zonal mean sections.

Tropical biases are largest in the stratocumulus and warm pool regions.

Bias in 850 hPa temperature (K) at day 3 in different models as indicated in the panel titles.



Structure of Tropical Precipitation Errors



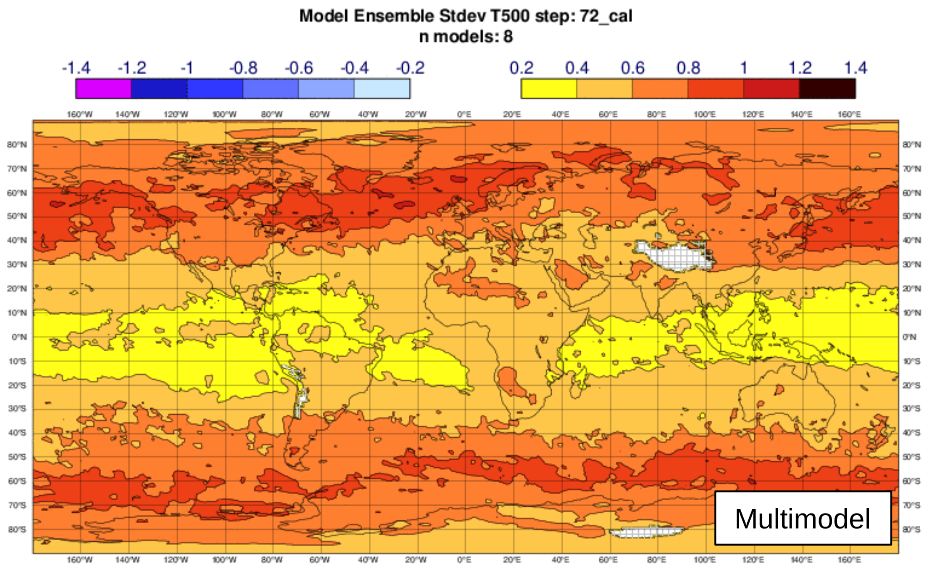
Average 10-day precipitation accumulation bias (mm) over southeast Asia and the Maritime Continent in models as indicated on the panel titles, computed w.r.t. GPM as shown at the top of the plot. Domain-averaged RMSE is annotated for each panel.

All models over-estimate precipitation over the warm pool and Maritime Continent.

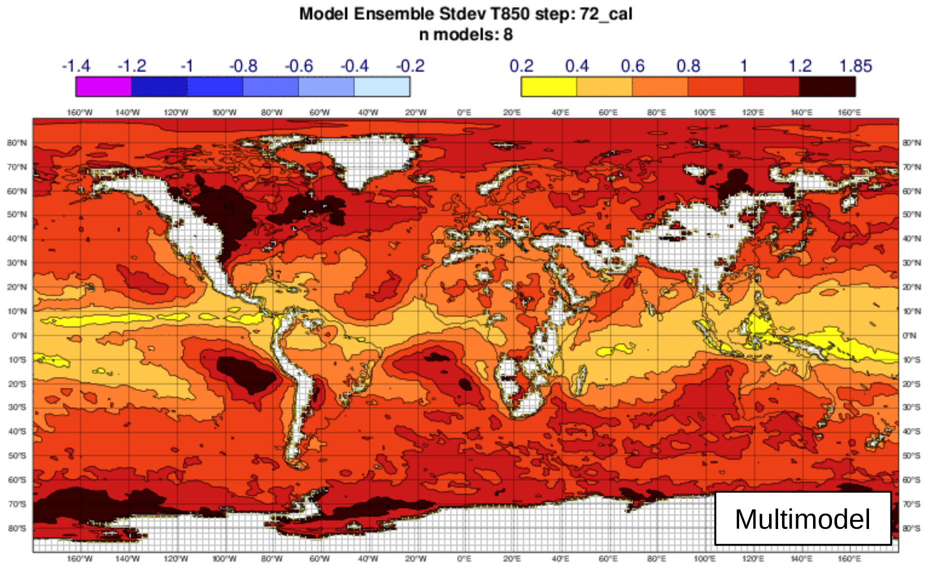
Most systems have negative biases over water, except ICON in which precipitation is reduced over land.

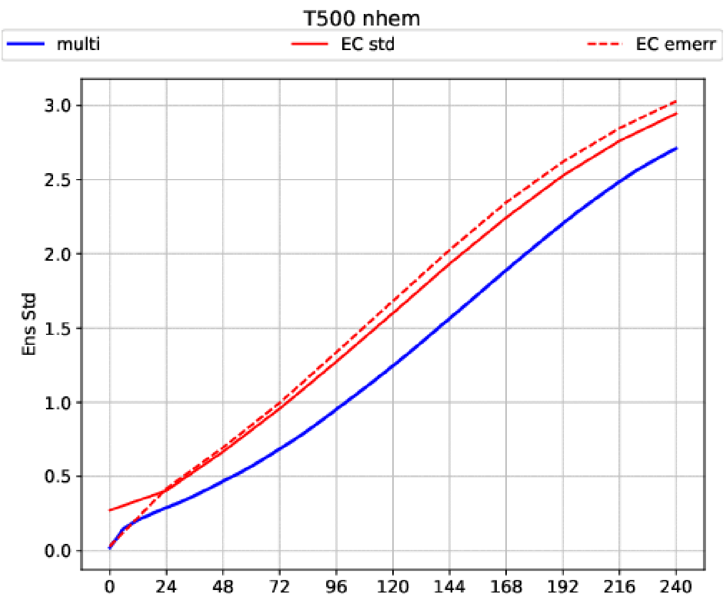
Divergence of Model Solutions: Temperature

T500



T850



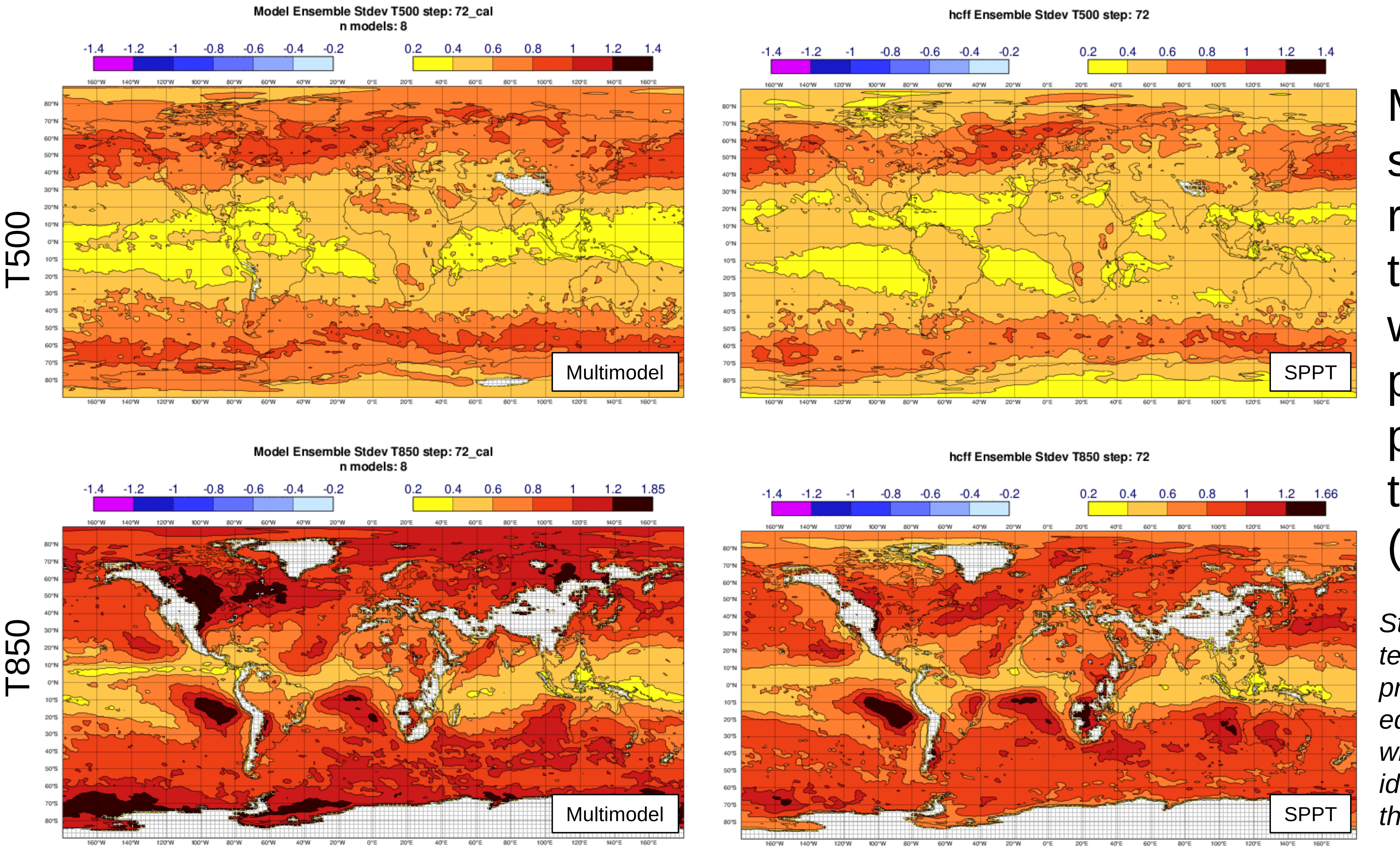


Multimodel ensemble spread in 500 hPa (top) and 850 hPa (bottom) temperature (K) at day-3 of forecast integrations. The evolution of this spread in the northern hemisphere (above) is shown for the multimodel (blue) and ECMWF (red dashed) systems, compared with ECMWF error standard deviation (red solid).

Model solutions diverge in the midlatitudes (aloft) an outside of the equatorial strip (PBL).

Normalization by climatological variability may help to identify regions where differences between the models are generating diversity.

Divergence of Model Solutions: SPPT Reference



Multimodel spread structure is remarkably similar to that achieved with stochastic perturbations to physics tendencies (SPPT).

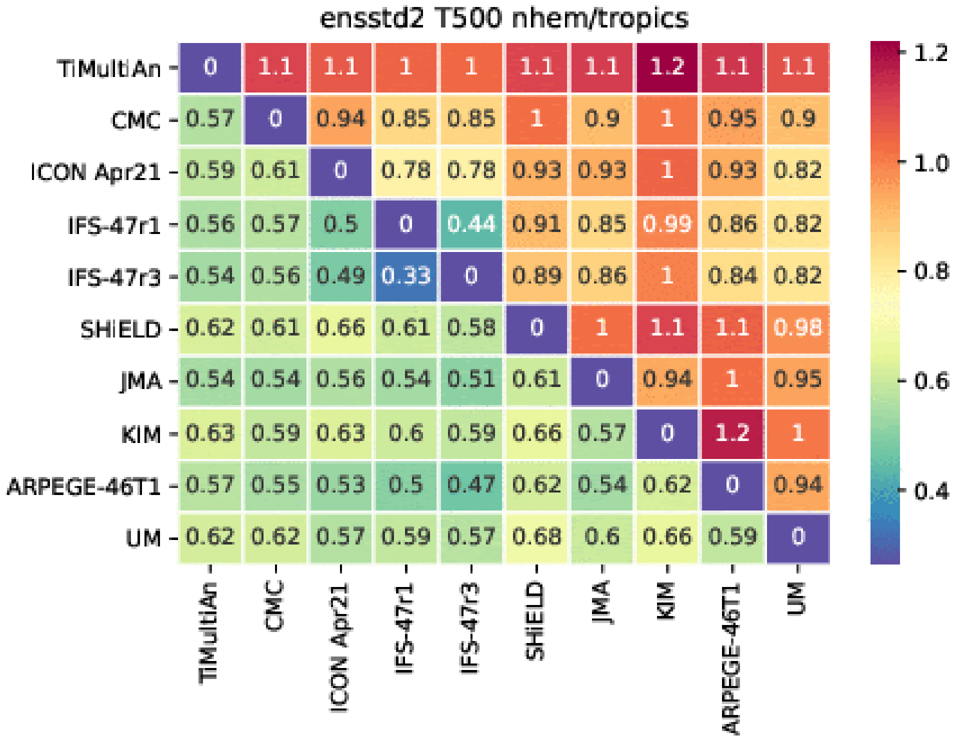
Standard deviation of temperature as shown on previous slide (left), and equivalent in an experiment with SPPT activated for identical initial conditions in the IFS (right).

Discussion

- Great engagement in the projects from many partners
- There is significant spin-up in most systems that appears to be related to moisture
- IFS in lead for most parameters followed by ICON, possibility a result of similarities between models (including Arpege)
- Results from additional diagnostic efforts:
  - Similar tropical cyclone track errors in IFS, ICON and SHiELD (Jan-Huey Chen, GFDL)
  - Most model suffer from a weak intensity bias and slow movement of extratropical cyclones (Duncan Ackerley, UKMO)
  - There appears to be a connection between peaks in individual forecast errors and representation of warm conveyor belts (Julian Quinting, KIT)
  - Initialization matters: all models with ECMWF ICs are better than SHiELD from GFS
- Comparison with model error representations may allow for further refinement of uncertainty estimates
- There is much to learn from this dataset: we have only scratched the surface so far!

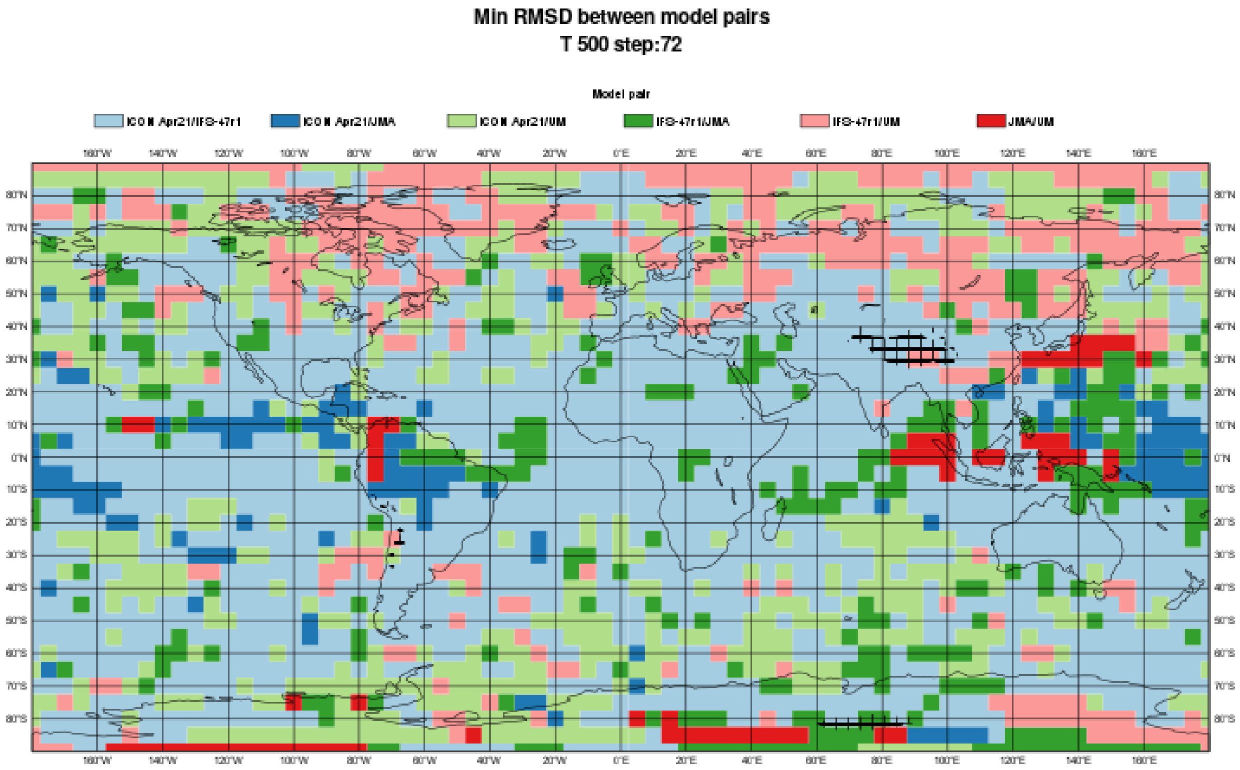
Further explorations

Forecast difference (RMS) between pairs of models, T500, day 3

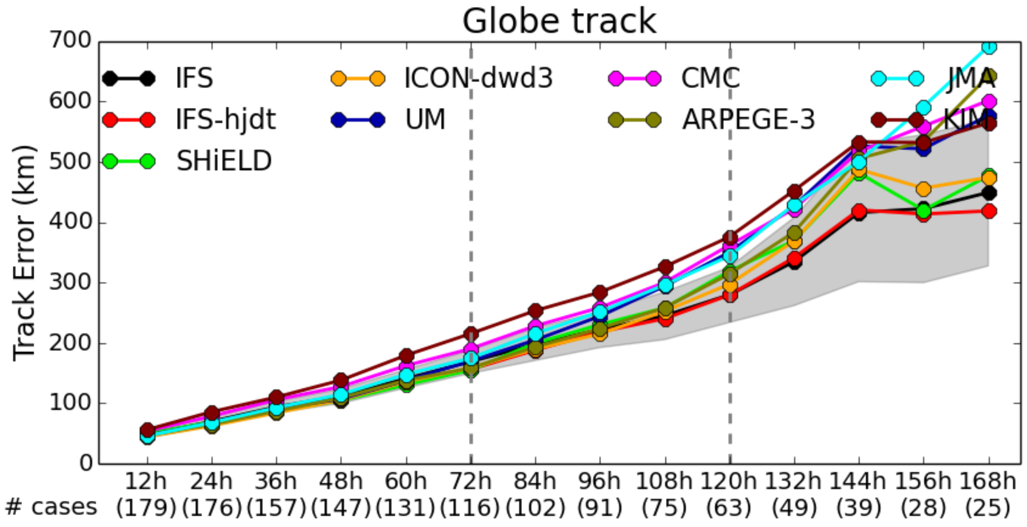


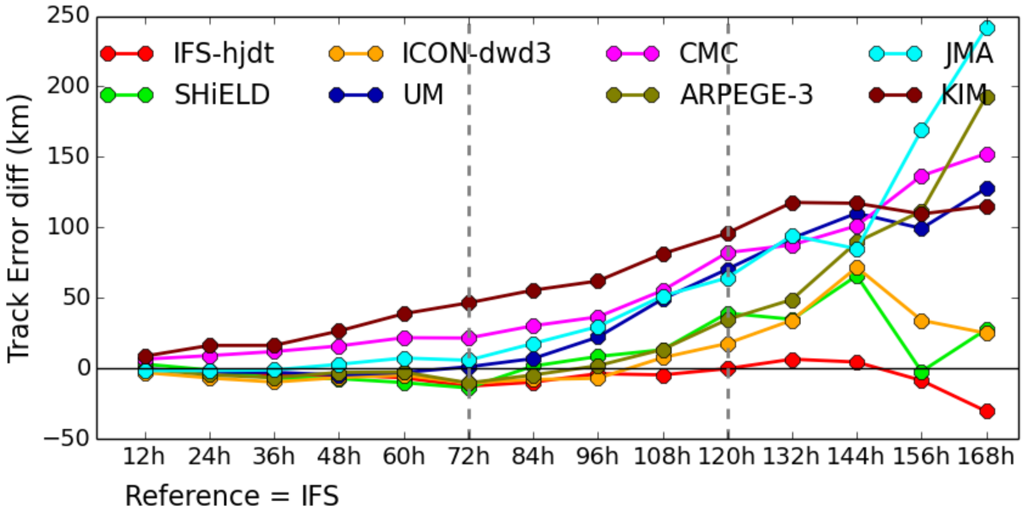
Matrix: Model pair RMS differences for T500, step 72 for N.Hem (top-left triangle) and Tropics (bottom-right).

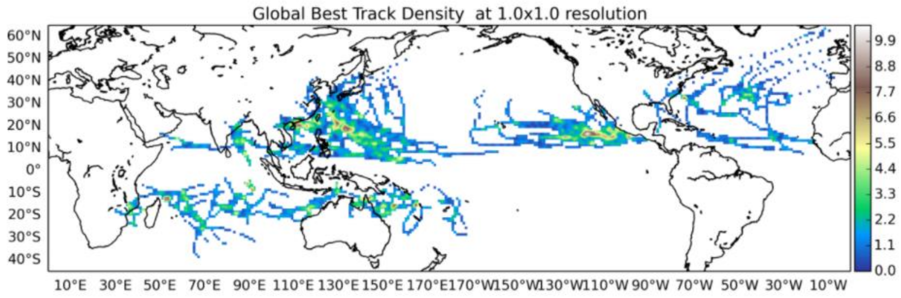
Map: Model pair with lowest RMSD in 5x5 degree boxes among ICON, IFS, UM and JMA.

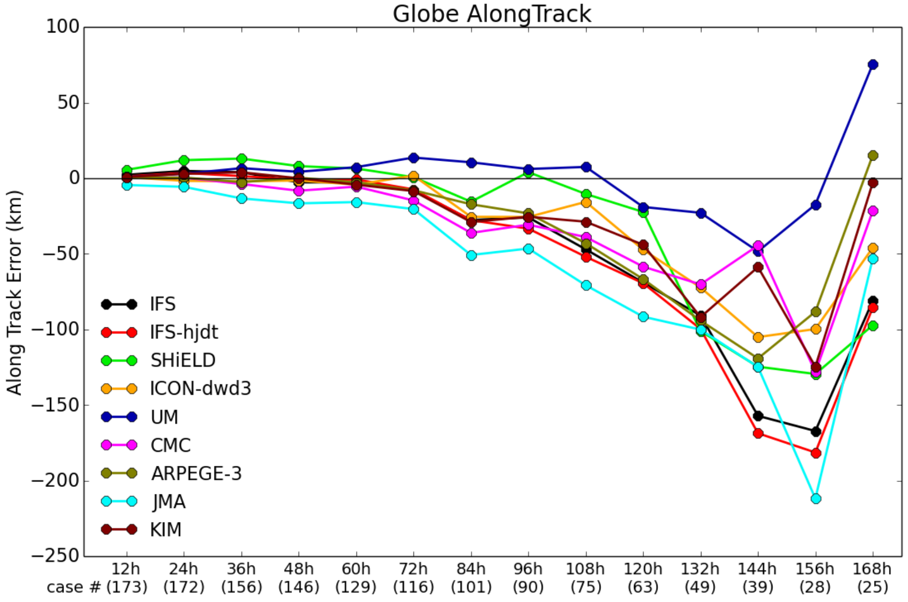


Tropical cyclone track errors (from Jan-Huey Chen, GFDL)







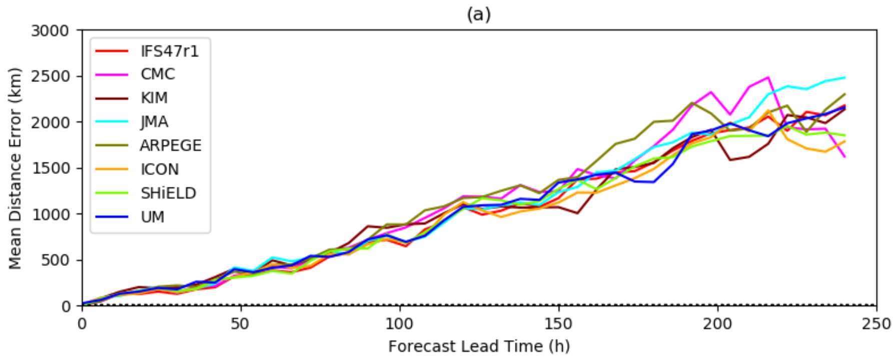


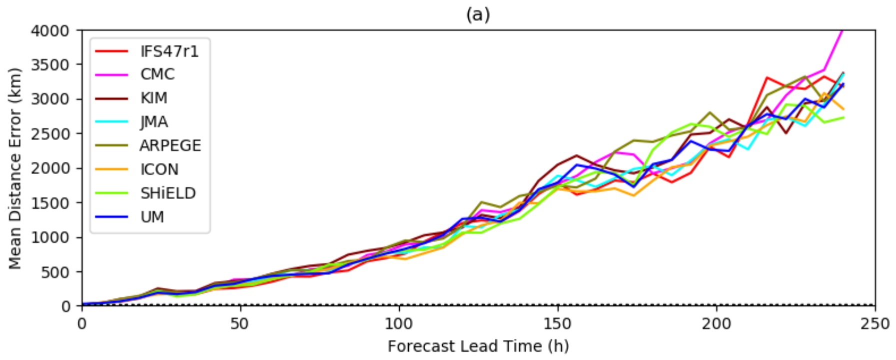
Extra-tropical cyclones (from Duncan Ackerley, UKMO)

Northern hemisphere

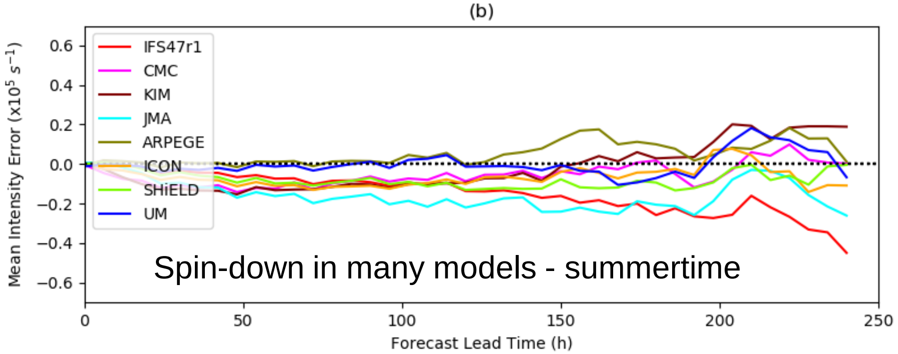
Southern hemisphere

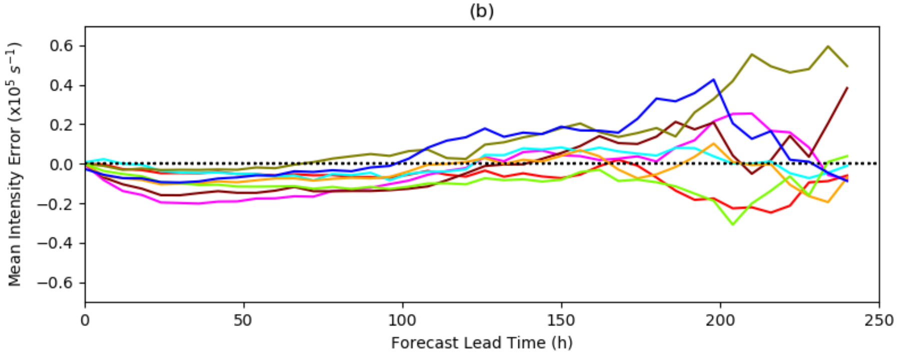
Position error



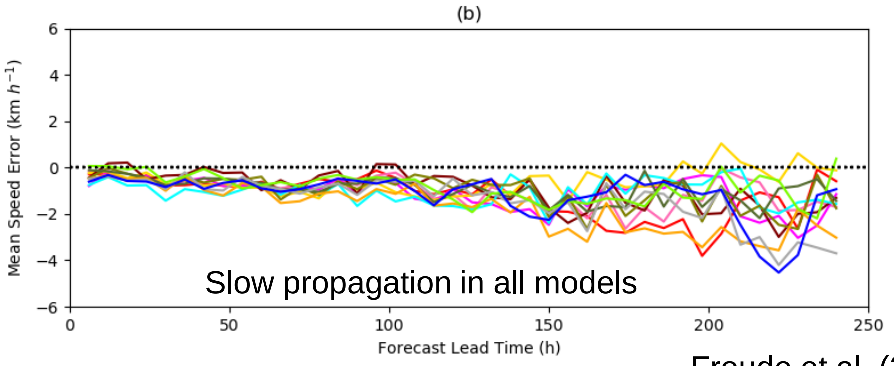


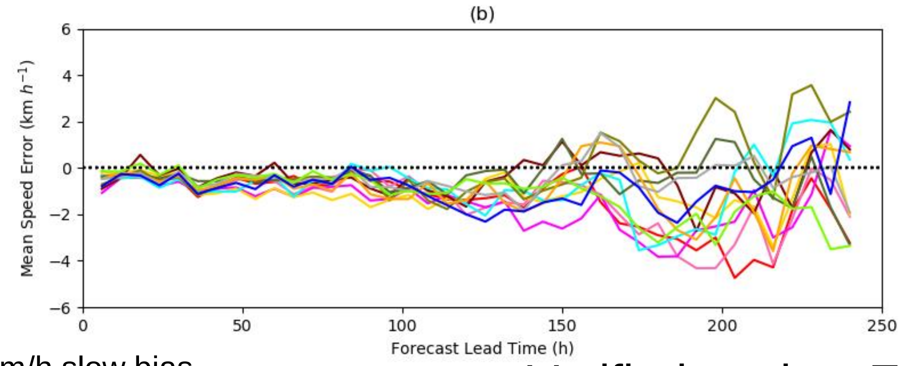
Intensity bias





Propagation speed



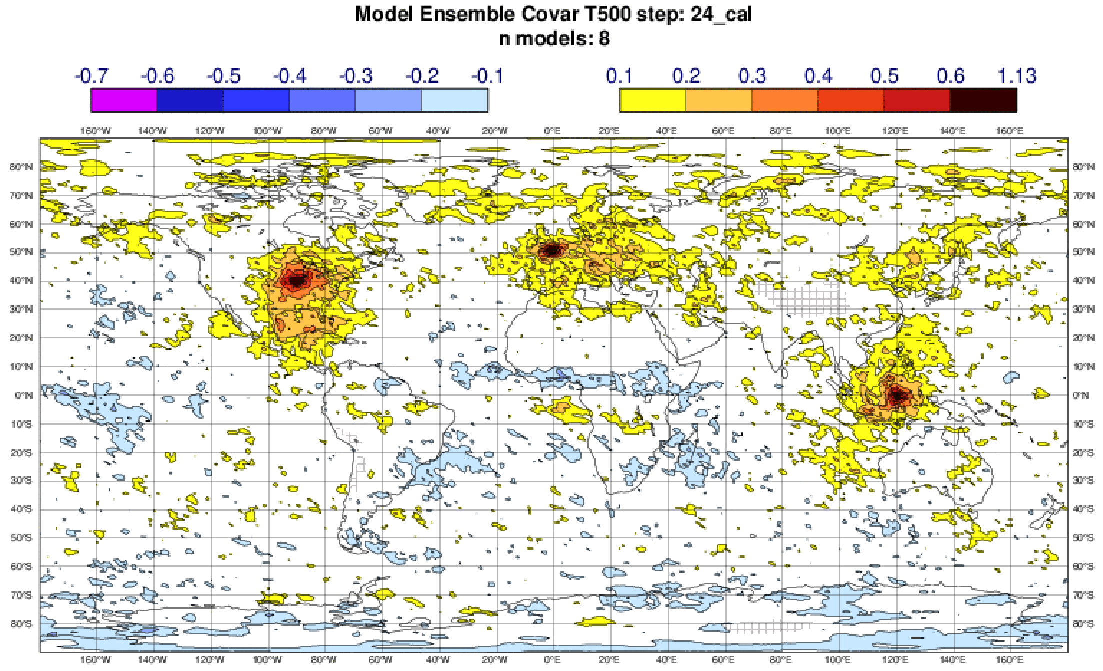


Froude et al. (2007): ~1 km/h slow bias

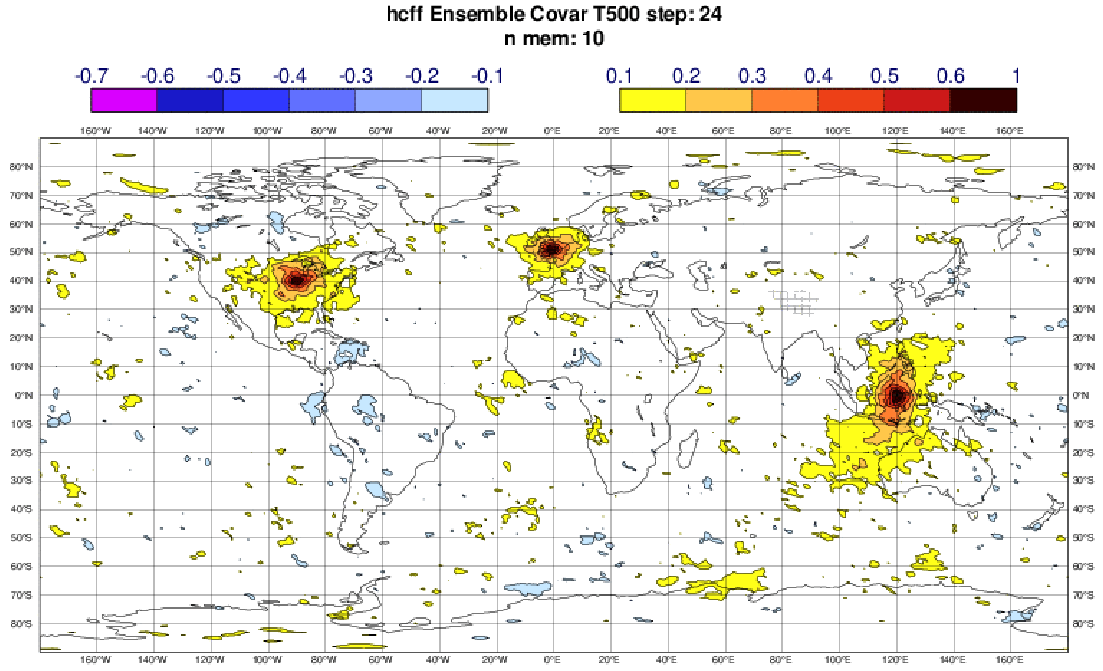
Verified against ECMWF analysis

Spatial ensemble correlations from 3 points – T500, step 24h

Multi-model ensemble

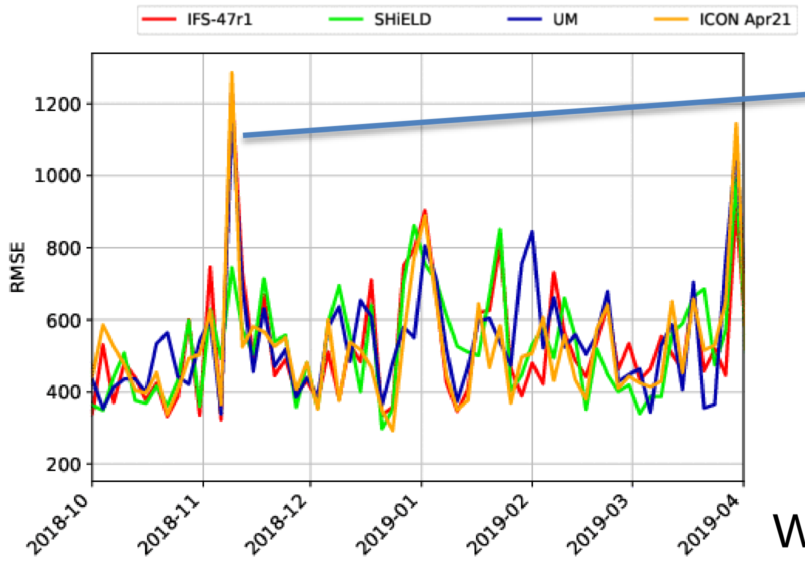


SPPT ensemble



Skill variability

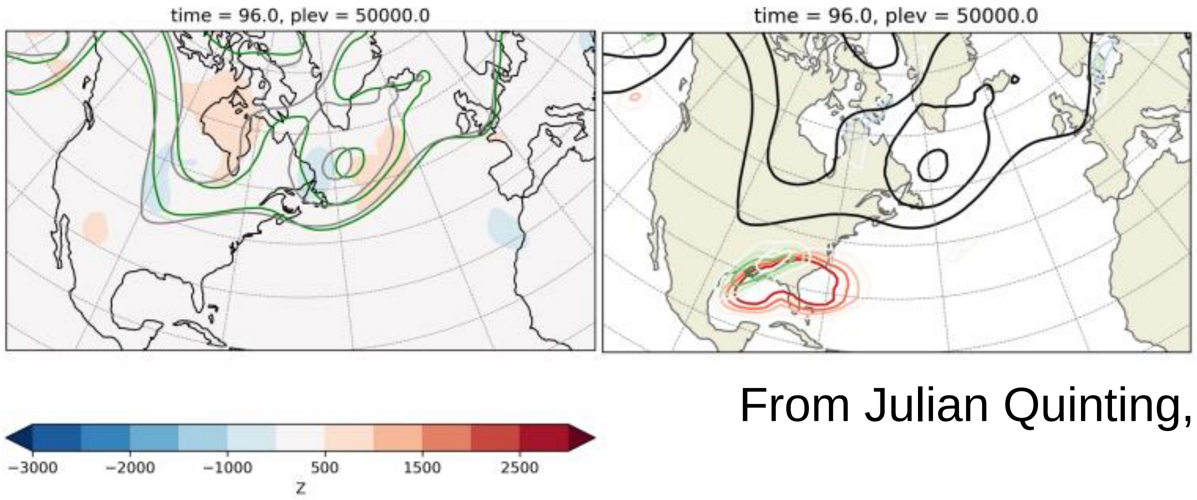
Z500, day 6 error over N. Atlantic



Shading: Z500fc – Z500an
Green contour: Z500an
Grey contour: Z500fc

Warm-conveyor belt diagnostics

Black contour: Z500fc
Red contour: WCB inflow prob.
FC Green contour: WCB acent prob.
FC Blue contour: WCB outflow prob. FC



From Julian Quinting, KIT